Proc. of the Intl. Conf. on Advances In Engineering And Technology - ICAET-2014

ISBN: 978-1-63248-028-6 doi: 10.15224/ 978-1-63248-028-6-03-60

Modelling of Reservoir Induced Earthquake Using Relevance Vector Machine

Shivam Mathur[1] and Kingshuk Misra[2]
[1] VIT University / School of Mechnical and Building Science, Vellore, India

[2] VIT University / School of Mechnical and Building Science, Vellore, India

***Abstract* - In large reservoirs, the water column alters in-situ stress state along an existing fault or fracture. The load of the water column is often so intense that it can significantly change the stress state leading to induced seismicity. Various empirical and calculative techniques are in place to predict the probable occurrence and magnitude of such seismic variations. This paper utilizes the Relevance Vector Machine (RVM) approach for prediction of Magnitude (M) of reservoir induced earthquake. RVM is developed in probabilistic framework. It produces sparse solution. RVM uses two input parameters namely depth of the reservoir and the other being a comprehensive parameter representing reservoir geometry for prediction of M.**

***Keywords* – reservoir–induced seismicity, relevance vector machine, radial basis function, Gaussian prior distribution, kernel function**

I. INTRODUCTION

Since the first instance of reservoir induced seismicity in Lake Mead in the early 1900's [1] reservoir induced seismicity (RIS) studies have been carried out in various parts of the world. When a dam is built and the reservoir filled with water the existing fault structure is subjected to varied stress and hence this often leads to movement along the fault plane giving rise to earthquakes [2].

The prediction of reservoir induced seismicity is important because it gives us the seismic potential of a dam after it has been built. Earlier artificial neural networks, statistical prediction model, fuzzy mathematics and gray system model have been used successfully for the prediction of reservoir induced seismicity [3]. The results from such models give us the magnitude (M) of the induced earthquake to varied degree of accuracy with the neural network method of prediction holding greater significance.

Radial basis functions (RBF) have also been utilized for the estimation of reservoir induced seismicity but although RBF networks are easy to train, when training is finished and it is being used it is usually slower [4].

In our study, we have used RVM to predict the magnitude of the earthquake induced by the reservoirs. RVM was introduced by Tipping [5] in 2001 and since then it has been used extensively in various civil engineering disciplines, namely, geotechnical engineering and water resource engineering [6-8]. RVM exploits a probabilistic Bayesian learning framework to derive an accurate prediction model. It achieves a sparse representation of the approximating function by structuring a Gaussian prior distribution. This is achieved by specifying independent Gaussian priors for each of the coefficients. Further comparison with conventional RBF and regressional analysis model is shown.

II. RELEVANCE VECTOR MACHINE

RVM produces sparse solutions using an improper hierarchical prior and optimizing over hyper parameters [9].

Let D={x_i, y_i, i=1, . . ., N} be a dataset of observed values.

where x_i = input, y_i = output, xi ϵ R^d and y_i ϵ R.

In this study, the input parameters are comprehensive parameter representing reservoir geometry (A) and height of the reservoir (B). Therefore, x = [A, B]. The output of the RVM model is magnitude of the earthquake induced by the reservoir. So, y = [M]. The output can be expressed as the sum of an approximation vector, z = $(z(x_1), \ldots, z(x_N))^T$, and zero mean random error (noise) vector, $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_N)^T$ where $\varepsilon_n \sim N(0,\sigma^2)$ and $N(0,\sigma^2)$ is the normal distribution with mean 0 and variance σ^2. So, the output can be expressed as:

$$y_n = z(x_{n,}\omega) + \varepsilon_n \quad (1)$$

where ω is the parameter vector.

Let us assume

$$p(y_n|x) \sim \boldsymbol{N}(z(x_n), \sigma^2) \quad (2)$$

where $\boldsymbol{N}(z(x_n), \sigma^2)$ is the normal distribution with mean $z(x_n)$ and variance σ^2. We can express $z(x)$ as a linearly weighted sum of m nonlinear fixed basis function, $\{\Phi_j(x)|j=1, \ldots, m\}$:

$$z(x,\omega) = \sum_{i=1}^{m} \omega_i \Phi_i = \Phi\omega \quad (3)$$

The likelihood of the complete data set can be written as:

$$p(y|w,\sigma^2) = (2\pi\sigma^2)^{-N/2}\exp\{-\frac{1}{2\sigma^2}||y - \Phi\omega||^2\} \quad (4)$$

where $y = (y_1, \ldots, y_N)^T$, $\omega = (\omega_0, \ldots, \omega_N)$ and

$$\Phi^T = \begin{bmatrix} 1 & K(x_1,x_1) & K(x_1,x_2) & \cdots & K(x_1,x_n) \\ 1 & \vdots & \vdots & \vdots & \vdots \\ 1 & K(x_1,x_1) & K(x_1,x_2) & \cdots & K(x_n,x_n) \end{bmatrix} \quad (5)$$

where $K(x_i, x_n)$ is a kernel function.

Proc. of the Intl. Conf. on Advances In Engineering And Technology - ICAET-2014

ISBN: 978-1-63248-028-6 doi: 10.15224/ 978-1-63248-028-6-03-60

In order to prevent over fitting, prior Automatic Relevance Detection (ARD) is set over weights.

$$p(w|a) = \prod_{i=0}^{N} N(\omega_i | 0, \alpha_i^{-1}) \qquad (6)$$

where α is a hyperparameter vector. It controls the deviation of each weight from zero [10]. Applying Bayes' rule, the posterior over all unknowns could be computed given the defined non-informative prior distribution:

$$p(w,\alpha,\sigma^2/y) = \frac{p(y/w,\alpha,\sigma^2)p(w,\alpha,\sigma)}{\int p(y/w,\alpha,\sigma^2)p(w,\alpha,\sigma^2)dwd\alpha d\sigma^2} \qquad (7)$$

It is hard to find full analytical solution of this integral. Thus, posterior distribution over the weights can be given by:

$$p(w/y,\alpha,\sigma^2) = \frac{p(y/w,\sigma^2)p(w/\alpha)}{p(y/\alpha,\sigma^2)} \qquad (8)$$

This results in a multi-variate Gaussian distribution resulting from the above mentioned posterior distribution:

$$p(w/y,\alpha,\sigma^2) = \boldsymbol{N}(\mu,\Sigma) \qquad (9)$$

where mean(μ) = $\sigma^{-2}\Sigma\Phi^T y$
and covariance (Σ) = $(\sigma^{-2}\Phi^T\Phi + X)$
with diagonal $X = \text{diag}(B_0, \ldots, B_N)$.

Maximizing the term $p(y/\alpha, \sigma^2)$ will give uniform hyperpriors over α and σ^2

$$p(w/y,\sigma^2) = \int p(y/w,\sigma^2)p(w/\alpha)dw = \left(\frac{(2\pi)^{-N/2}}{\sqrt{|\sigma^2 + |\Phi X^{-1}\Phi^T}}\right) \times exp\left\{-\frac{1}{2}z^T(\sigma^2 + \Phi X^{-1}\Phi^T)^{-1}z\right\} \qquad (10)$$

This maximizing term is known as "evidence for hyperparameter"[11]. Outcome of this optimization is that many elements of α will become infinite and therefore, there will be very few non-zero weights. These weights are considered as relevance vectors. Now, the equation of RVM will finally be:

$$f(x) = \sum_{i=0}^{nrv} w_i K(xx_i) \qquad (11)$$

where *nrv* are number of relevance vectors.

In our study, RVM methodology has been used to calculate magnitude of the reservoir induced seismicity. The magnitude of the earthquake (M) can be determined by:

$$M = 0.995E + 1.44 \qquad (12)$$

The comprehensive parameter (E) has been defined as:

$$E = \frac{SH}{V} \qquad (13)$$

where S, H and V are reservoir surface area, maximum reservoir depth and reservoir capacity respectively.

III. APPLICATION

RVM makes use of Gaussian Kernel function. It requires datasets to make a model and then test its efficiency. During formulation the datasets were divided into two sets:

A. *Training datasets*

These are required for the construction of RVM model. For this study, we have used 24 datasets for the construction of the model. Table 1 shows the normalised datasets used to carry out the method.

B. *Testing datasets*

This is used to verify the developed RVM. The remaining 6 datasets were used as testing datasets.

All the datasets were normalised between 0 and 1.

TABLE 1
NORMALISED DATASETS USED FOR RVM

Height (H)	Comprehensive Parameter (E)	Magnitude (M)
0.162297	0.475655	0.325581
0.208073	0.258427	0.325581
0.362047	0.47191	0.604651
0.253849	0.494382	0.674419
0.607574	0.441948	0.674419
0.320433	1	1
0.570121	0.779026	0.325581
0.037453	0.089888	0.395349
0.31211	0.988764	0.953488
0.586767	0.505618	0.651163
0.212235	0	0.465116
0.06367	0.606742	0.837209
1	0.003745	0.55814
0.661673	0.378277	0.837209
0.145651	0.217228	0.604651
0.449438	0.498127	0.209302
0.403662	0.389513	0.325581
0.077403	0.179775	0.209302
0.245526	0.707865	0.767442
0.216396	0.516854	0.930233
0.195589	0.40824	0.604651
0.151061	0.269663	0
0.349563	0.265918	0.162791
0.020807	0.265918	0.209302

Proc. of the Intl. Conf. on Advances In Engineering And Technology - ICAET-2014

ISBN: 978-1-63248-028-6 doi: 10.15224/ 978-1-63248-028-6-03-60

0.070745	0.539326	0.255814
0.122347	0.599251	0.55814
0	0.370787	0.162791
0.022056	0.539326	0.627907
0.258011	0.123596	0.27907
0.258011	0.123596	0.255814

IV. RESULTS AND DISCUSSION

The final equation constructed by the RVM in terms of weights (w) and width (σ) with the help of 24 training datasets to calculate M is as follows:

$$M = \sum_{I=1}^{24} w_i e^{\frac{-(x_i - x)(x_i - x)^T}{2\sigma^2}} \qquad (14)$$

The Fig. 1 shows the relation between the original training datasets and RVM predicted datasets. As it is shown the predicted datasets are identical to the original datasets with a good coefficient of correlation (R). Coefficient of correlation (R) is the most important criteria to determine the performance of RVM. Fig. 2 shows the relation between original testing datasets and RVM predicted datasets. The identical values validate the performance of the RVM. Another important factor to be discussed is the weights of the training datasets. In Table 2, we can see that variation of the values is only marginal. It shows the sparseness of the RVM. This shows that the performance function will be efficient and smooth.

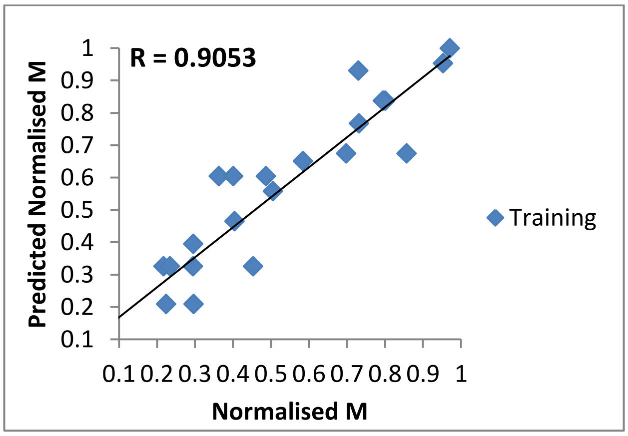

Fig.1 Performance of training datasets

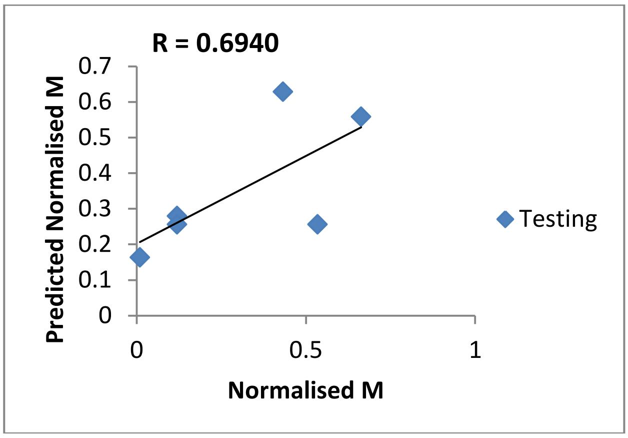

Fig. 2 Performance of testing datasets

TABLE 2
VALUE OF WEIGHTS

Height (H)	Comprehensive Parameter (E)	Weights (w)
84	3.18	0
95	2.6	0
132	3.17	0.167
106	3.23	0.741
191	3.09	0
122	4.58	0.9706
182	3.99	0.2342
54	2.15	0.2628
120	4.55	0
186	3.26	0.4913
96	1.91	0.3954
60.3	3.53	0.7826
285.3	1.92	0.5051
204	2.92	0.7344
80	2.49	0.3393
153	3.24	0
142	2.95	0.1833
63.6	2.39	0
104	3.8	0.7098
97	3.29	0
92	3	0.1169
81.3	2.63	0
129	2.62	0
50	2.62	0

But the marginal difference between the original testing values and RVM predicted testing values is due to over training of the performance function. Therefore, to get better performance of the testing datasets, over training should be avoided.

ISBN: 978-1-63248-028-6 doi: 10.15224/ 978-1-63248-028-6-03-60

In Fig. 3, the Coefficient of correlation (R) of RVM is compared with that of the conventional statistical approach, i.e., regression analysis and that of RBF.

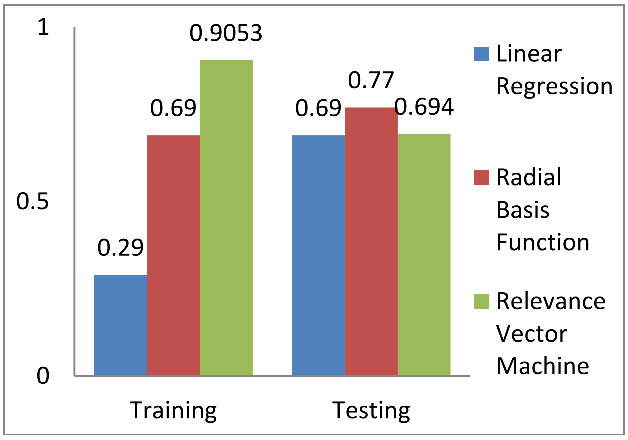

Fig. 3 Coefficient of correlation

V. CONCLUSION

This study shows the efficient RVM for determination of magnitude of the earthquake generated due to the stresses developed in the reservoir. The final result has been compared with the linear regression model and RBF. The similar results by both RVM and RBF suggest that RVM is equally acceptable than RBF. The proposed RVM is not a replacement of RBF but just an alternative.

AKNOWLEDGEMENT

The authors wish to thank Dr. Pijush Samui for his unconditional support and encouragement for this study.

REFERENCES

[1] Pradeep Talwani, "On the Nature of Reservoir-induced Seismicity," *Pure appl. Geophys*, 1997, pp 473–492.

[2] Pradeep Talwani and Steve Acree, "Pore pressure diffusion and the mechanism of reservoir-induced seismicity," *Pure and applied geophysics*, Volume 122, Issue 6, 1984/85, pp 947-96.

[3] Jian Deng, Desheng Gu, Xibing Li and Zhong Qi Yue, "Structural reliability analysis for implicit performance functions using artificial neural network," *Structural Safety* Vol. 27, 2005, pp. 25–48.

[4] Ghassem Habibagahi, "Reservoir induced earthquakes analyzed via radial basis function networks," *Soil Dynamics and Earthquake Engineering*, Volume 17, 1998, 53-56.

[5] M.E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J.Mach. Learn*., Vol. 1, pp. 211–244, 2001.

[6] Samui, P., "Relevance vector machine applied to settlement of shallow foundation on cohesionless soils," *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, Vol. 2, No.1, pp.41 – 47, 2008.

[7] Samui, P., "Seismic liquefaction potential assessment by using Relevance Vector Machine," *Journal Earthquake Engineering and Engineering Vibration.* Vol. 6, No.4, pp. 331-336, 2007.

[8] Ghosh, S. and Mujumdar, P. P., "Statistical downscaling of GCM simulations to streamflow using relevance vector machine," *Advances in Water Resources*, Vol. 31, pp.132–146, 2008.

[9] Pijush Samui, Tim Lansivaara and Dookie Kim, "Utilization relevance vector machine for slope reliability analysis," *Applied Soft Computing*, Vol. 11, pp. 4036-4040, 2011.

[10] B. Scholkopf and A.J. Smola, "Learning with Kernels: Support Vector Machines," Regularization, Optimization and Beyond, *MIT Press*, Cambridge, Massachusetts, 2002.

[11] D.J. MacKay, "Bayesian Methods for Adaptive Models," Ph.D. Thesis, *Department of Computational and Neural Systems*, California Institute of Technology, Pasadena, California, 1992.